Sequence Listing

Individual Applicant

Street :
City : Athens
State : GA
Country : USA
PostalCode : 30602-7272
PhoneNumber : 706-542-0928
FaxNumber : ___-___-____
EmailAddress : wparrott@uga.edu
<110> LastName : Parrott
<110> FirstName : Wayne
<110> MiddleInitial : A
<110> Suffix : PhD.

Individual Applicant

Street :
City : Athens
State : GA
Country : USA
PostalCode : 30602-7272
PhoneNumber : 706-542-0928
FaxNumber : ___-___-____
EmailAddress :
<110> LastName : LaFayette
<110> FirstName : Peter
<110> MiddleInitial : R
<110> Suffix : PhD.

Individual Applicant

Street : 1907 South Milledge Ave
City : Athens
State : GA
Country : USA
PostalCode : 30605-____
PhoneNumber : ___-___-____
FaxNumber : ___-___-____
EmailAddress :
<110> LastName : Kane
<110> FirstName : Patrick
<110> MiddleInitial : M
<110> Suffix :

Application Project

<120> Title : Arabitol or Ribitol As Positive Selectable Markers
<130> AppFileReference : UGA-855R
<140> CurrentAppNumber :
<141> CurrentFilingDate : ____-__-__

Sequence

<213> OrganismName : Escherichia coli
<400> PreSequenceString :
atgaacgaac aatttacatg gctgcacatc gggttaggtt cttttcatcg cgcacatcag 60
gcgtggtatc tacaccgttt gcaggtgatg ggcgataaac gctggagcat tgctgcgggc 120
aatattcgta atgatgctga acatgtcgta caggcactca gtgcacagaa aggtcgctat 180
gtgctggaaa ccgtcagccc ggaaggggta agcgaatatg aagagatcac ctcaattcag 240

aagttgatac cgtggcaggc agatttacaa ccgctgattg ctgaaggggc agatccgaag 300
acaaaagtga ttgctttcac cgtcaccgaa ggcgggtact acctgaatac cagtcacaaa 360
ctggaagtta acaatcctga tttagcggca gatcttaaag ggggatgcaa aacaatttac 420
ggtgttatta cccgtatcct cgaagcgcgt atggcaaata acgccggacc actaaccctg 480
atgaacgaac aatttacatg gctgcacatc gggttaggtt cttttcatcg cgcacatcag 540
gcgtggtatc tacaccgttt gcaggtgatg ggcgataaac gctggagcat tgctgcgggc 600
aatattcgta atgatgctga acatgtcgta caggcactca gtgcacagaa aggtcgctat 660
gtgctggaaa ccgtcagccc ggaaggggta agcgaatatg aagagatcac ctcaattcag 720
aagttgatac cgtggcaggc agatttacaa ccgctgattg ctgaaggggc agatccgaag 780
acaaaagtga ttgctttcac cgtcaccgaa ggcgggtact acctgaatac cagtcacaaa 840
ctggaagtta acaatcctga tttagcggca gatcttaaag ggggatgcaa aacaatttac 900
ggtgttatta cccgtatcct cgaagcgcgt atggcaaata acgccggacc actaaccctg 960
ctgaattgcg ataacgtgcg ccataatggt gaacgtttcc atgatggcct ggttgagttt 1020
ctccagctaa ctggcaaaca ggatgtcatc gactggctga gtacaaatac cacttgcccg 1080
aataccatgg ttgaccgcat tacgcctcgt ccggcagcag aacttccggc acggatcaag 1140
gctcaaacgg gtattgccga taaagcgccg gtaatgggcg aaacctttat ccagtgggtc 1200
gtggaagata atttccgtga tgtccgtccg gcactggaga aggtcggtgt cgaactggtg 1260
gcgtcggtaa tcccctatga agaggcgaag attcgcattc ttaactcttc acacagttgc 1320
atcgcctggg caggtacgtt aatcggtcaa aaatatatcc acgaaagcac aatgaccgat 1380
tttatctatc agattgccga ccgctacgtg acggaagatg tcattccttg cttgggcgat 1440
aacggtatcg atttgccaac ctaccgtgat gttgtactca agcgttttac caatccacat 1500
attcaggaca ccaaccaacg cgtcgctgcg gatggtttct cgaaaattcc ggcgatgatt 1560
gcccccacac tgcgagagtg ctaccagcga ggcgttcgcc cgaatgccac cgccatgtta 1620
cctgcactgt tttacgtatt catggagcag tggcatcacg gcaaactgcc ctatgaatat 1680
caggatggca tccttgatgc accagctgtc catgcaatgt tacagtctgc cgatcccgtc 1740
gctgtttatg ccagtgataa agcgctgttt ggcgatttaa ccgaacgtga agattttgcc 1800
gcgttgttgc gcgaaaaaat cgctgacgtc tacgcgttaa ttaactaa 1848

<212> Type : DNA
<211> Length : 1848
SequenceName : SEQ ID NO 1
SequenceDescription : arabitol dehydrogenase

Custom Codon

Sequence Name : SEQ ID NO 1

Sequence

<213> OrganismName : Escherichia coli
<400> PreSequenceString :
atcgattgag cagtttgctt cacacggcag ctaaattccc gttcagtgcg tgcaagcaaa 60
cagagactat aaattcgccc tggtaaaagg attatatgat gaatcactct gtgccctcta 120
tgaatactcc ccttaatggc aaagttgcag ctatcactgg cgctgcgtca ggtattggcc 180
tgcaatgtgc aaaaacgctg ctcgatgcag gagcaaaggt agtattgatt gaccgggaag 240
gcgacaaact gcacaagatt gtcgctgagt taggcgaaaa cgcgtacgcg ctgcaactcg 300
atctcttcaa taatcagcaa gtcgataaca tgctggcgga cattatcgaa ctggcgggtg 360
ggctggatat ttttcatgcc aatgcaggcg cttatattgg cggcccagtg gctgaaggtg 420
atccagatgt ctgggatcgt gtgttaaatc tgaatataaa tgcggcgttt cgctgtgtcc 480
gtgcagtcct gccgcatatg attgcgcaga ggtcgggcga tataatttct accagttcca 540
tcgcgggcgt cgttccggtt atctgggaac cgatctacac cgcgtccaaa tttgccgttc 600
aggcattcgt acacactacc cgccgccagg tttctcaata tggcgtgcgt gtgggtgcgg 660
tgctgccagg accagtagtc actgccctgc ttgatgactg gccaaaagcc aaaatggaag 720
aagccctggc aaatggtagc ctgatgcaac cgattgaagt ggcggaatca gtattgttta 780
tggtgacccg ctcgaaaaat gtcaccgtgc gagatttagt gatcctgcct ggcagtgtcg 840
atctgtaagg gcgcaatcat gacaataacc aaaaccgtta ttggtgttga tgtgggatca 900
ggcagtgtcc gcgccgggat ttttgatctc aacggatctc tgctatccca tgccacagaa 960
aaaatcacga ctacgcggcg cagcggaagc cgcgtggaac agtccagcca ggagatctgg 1020
caggcggtct gttcatgtat tcgtaatgcg ctcactctgg cagacgtttg tgcacaaagt 1080
gtggcaggca tcggttttga tgccacctgt tctctggtgg tactggataa aaacggtgat 1140
ccattgcctg tcagcccgga aggagatgca aagcaaaata tcattgtgtg gatggatcac 1200
cgcgccaccg aacaagcaga gcgaatcaat gccactcacc atccggtgct gaactacgtc 1260
ggtggtaaaa tttcgcctga aatggaaaca ccgaaaattc tctggctgaa agaaaatatg 1320
ccagagatct acgaacgtgc cggacaattt ttcgatctgg ccgattttct gacctggcgg 1380
gctaccggtg atttagcgcg ttcagtatgc actgttacct gtaaatggac gtggctggca 1440

```
catgaaaatc gctgggatcc agattatttc cgcaccattg gccttgcaga gttagcggat    1500
gaagatttta ttcgcattgg tcatcatatt gtttctcccg gaacaccttg tggaaatggt    1560
ttaacagcac aagccgcggc agagatggga ttactccccg gcacacctgt cgctgtaggg    1620
ttaattgatg ctcacgctgg tggcatcggt acggttggcg tagaaggtgg agcgctgaac    1680
aatctcgcgt atgttttcgg cacttcttca tgcaccatgg catctaccac ttctccctcg    1740
tttgtaccgg gtgtctgggg gccgtattac agtgcgatgg ttccaggtct gtggttagtt    1800
gaaggcgggc aaagtgctgc gggagcagct attgaccagc tacttgattt ccatccggct    1860
gttgaagaag ctcgcgaaat ggcacaacgt gtgaatcagc ccctccccgt ctggcttgct    1920
gatcgaatcc tcgaaaaaac ggcgcaacca tcagatgctg tcgccctggc gaaagggcta    1980
cacgtggtgc cggaatttct cggaaatcgc gcccccttcg cagatcctca tgccagagcg    2040
gtaatttgtg gcctgggtat ggagcgagat ctggataatt tactcgcctt gtatatcgct    2100
ggattatgcg gaattggtta tggtctgcgc caaattctcg acgctcaaac agcgcaggga    2160
gtagtgagta aaaatatcgt tattagcggc ggtgccgggc agcatccact ggtacgacaa    2220
attctggcag atacctgcgg tattccggtc attaccacgc aatgctgcga accggtttta    2280
ttaggctcgg ccattcttgg tgctgtcgcc ggaaatattg caccttctgt tggcgaagcg    2340
atgcaacaat tcacccatgt ggataaatat tattatccgc aagaacgcta tcaatctctt    2400
catcatcgtc gatatgaggc ctataagcag ttgcagcata ctgcaaaatt actcagagac    2460
taattaacca gccgcctgac gctgttttca ggcaatcact aataacgact cactccggta    2520
atatcccgga gtgcattcat ctgcacccta aaaacgaggt ttatatgtcc agaaataata    2580
aacagtggtt gggtttgcca ctgcatctga tatggggata tatcgccatc gcagtattta    2640
tgactggtga tggtttcgaa ctcgcattct tatcgcacta tattaaatcg ttaggcttca    2700
cacctgcgga agcctctttt gcctttacgc tctacggcct ggcggctgcc ctttccgcct    2760
gggtttctgg ggtagtagcg gaaatcatca cgccgcaaaa aaccatgctg attggttttg    2820
tcctatggtg cgtattccat gttctgtttc tggtctttgg attaggacag gcaaactatg    2880
ggttaatcct gctgttttac gggattcgtg gtttagcgta tccgctattt ctctactcat    2940
ttatcgtcgt tattattcat aacgtgcgta gcgaaaattc cagttctgct ctgggttggt    3000
actgggcggt atattcagta gggatcggtg ttgctggcag ttatattccc agttttacga    3060
taccgattat gggtgaaatg ggaaccttat ggctggcact ggcgttctgc tttgctggcg    3120
gtgtcattgc catgatctcc ttgcgtcatg ttaaaacgcc tggacatatg cataatttaa    3180
ctccccgtga gaagtttgca gaattaagtc gggcagtaac tttactttat accaaccgca    3240
atatttttct ctccagtatt gtgcgcatta taaatacctt atcgttattc ggttttgcgg    3300
tcattatgcc aatgatgttt gtggatgaac tgggattcac cacctctgaa tggttgcagg    3360
tctgggcggc atttttcttt accactattt tctctaatat tttttggggg attgtggcag    3420
aaaaaatggg ctggatgcgt gttattcgct ggtttggttg cctcgggatg gcagcatcaa    3480
gtttagcgtt ttactacatg ccgcaatact ttggtcacaa ctactggatg gcaatgattc    3540
cggcgattgc tctgggaact tttgttgctg catttgtgcc gatggccgct gtcttcccgg    3600
cactggaacc aaaacacaaa ggtgctgcaa tctcggttta caacctctct gcgggtatgt    3660
ctaacttcct ggctccggca attgccgtgg tgttattacc gtggtttagc actatcggtg    3720
tggtcattgc ctatacagca ttgtatctat tggcctttgt cctttgcgca ttcattcgcg    3780
ttgagcagcc aggattcagt tctgcgccag tgactgagaa ggcattgaat atctcctgaa    3840
aaacgaaacg catcaggcac tcatcctctc cctcatggga gaggatgatt tcacatcagg    3900
caatagtgac tttgttatcg agataaacgt cctgcacggc gttaatcagt ttcacgccgt    3960
cagccatcga t                                                        3971
```

<212> Type : DNA
<211> Length : 3971
SequenceName : SEQ ID NO 2
SequenceDescription : ribitol operon

Feature

Sequence: SEQ ID NO 2:
<221> FeatureKey : misc_feature
<222> LocationFrom : 96
<222> LocationTo : 848
Other Information : ribitol dehydrogenase coding region
CDSJoin : No

Feature

Sequence: SEQ ID NO 2:
<221> FeatureKey : misc_feature
<222> LocationFrom : 859
<222> LocationTo : 2463
Other Information : ribitol kinase coding region

CDSJoin : No

Feature

Sequence: SEQ ID NO 2:
<221> FeatureKey : misc_feature
<222> LocationFrom : 2565
<222> LocationTo : 3839
Other Information : ribitol transporter coding region
CDSJoin : No

Custom Codon

Sequence Name : ribitol operon

Sequence

<213> OrganismName : Escherichia coli
<400> PreSequenceString :
MMNHSVPSMN TPLNGKVAAI TGAASGIGLQ CAKTLLDAGA KVVLIDREGD KLHKIVAELG 60
ENAYALQLDL FNNQQVDNML ADIIELAGGL DIFHANAGAY IGGPVAEGDP DVWDRVLNLN 120
INAAFRCVRA VLPHMIAQRS GDIIFTSSIA GVVPVIWEPI YTASKFAVQA FVHTTRRQVS 180
QYGVRVGAVL PGPVVTALLD DWPKAKMEEA LANGSLMQPI EVAESVLFMV TRSKNVTVRD 240
LVILPGSVDL 250
<212> Type : PRT
<211> Length : 250
SequenceName : SEQ ID NO 3
SequenceDescription : ribitol dehydrogenase

Sequence

<213> OrganismName : Escherichia coli
<400> PreSequenceString :
MTITKTVIGV DVGSGSVRAG IFDLNGSLLS HATEKITTTR RSGSRVEQSS QEIWQAVCSC 60
IRNALTLADV CAQSVAGIGF DATCSLVVLD KNGDPLPVSP EGDAKQNIIV WMDHRATEQA 120
ERINATHHPV LNYVGGKISP EMETPKILWL KENMPEIYER AGQFFDLADF LTWRATGDLA 180
RSVCTVTCKW TWLAHENRWD PDYFRTIGLA ELADEDFIRI GHHIVSPGTP CGNGLTAQAA 240
AEMGLLPGTP VAVGLIDAHA GGIGTVGVEG GALNNLAYVF GTSSCTMAST TSPSFVPGVW 300
GPYYSAMVPG LWLVEGGQSA AGAAIDQLLD FHPAVEEARE MAQRVNQPLP VWLADRILEK 360
TAQPSDAVAL AKGLHVVPEF LGNRAPFADP HARAVICGLG MERDLDNLLA LYIAGLCGIG 420
YGLRQILDAQ TAQGVVSKNI VISGGAGQHP LVRQILADTC GIPVITTQCC EPVLLGSAIL 480
GAVAGNIAPS VGEAMQQFTH VDKYYYPQER YQSLHHRRYE AYKQLQHTAK LLRD 534
<212> Type : PRT
<211> Length : 534
SequenceName : SEQ ID NO 4
SequenceDescription : ribitol kinase

Sequence

<213> OrganismName : Escherichia coli
<400> PreSequenceString :
MSRNNKQWLG LPLHLIWGYI AIAVFMTGDG FELAFLSHYI KSLGFTPAEA SFAFTLYGLA 60
AALSAWVSGV VAEIITPQKT MLIGFVLWCV FHVLFLVFGL GQANYGLILL FYGIRGLAYP 120
LFLYSFIVVI IHNVRSENSS SALGWYWAVY SVGIGVAGSY IPSFTIPIMG EMGTLWLALA 180
FCFAGGVIAM ISLRHVKTPG HMHNLTPREK FAELSRAVTL LYTNRNIFLS SIVRIINTLS 240
LFGFAVIMPM MFVDELGFTT SEWLQVWAAF FFTTIFSNIF WGIVAEKMGW MRVIRWFGCL 300
GMAASSLAFY YMPQYFGHNY WMAMIPAIAL GTFVAAFVPM AAVFPALEPK HKGAAISVYN 360
LSAGMSNFLA PAIAVVLLPW FSTIGVVIAY TALYLLAFVL CAFIRVEQPG FSSAPVTEKA 420
LNIS 424
<212> Type : PRT
<211> Length : 424
SequenceName : SEQ ID NO 5
SequenceDescription : ribitol transporter